Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 128 (2015) 44 – 53

Procedia Engineering

www.elsevier.com/locate/procedia

3rd European STAMP Workshop, STAMP EU 2015

The Risk Situation Awareness Provision Capability and its degradation in the Überlingen accident over time

Maria Mikela Chatzimichailidou[a,*], Ioannis M. Dokas[a]

[a]*Democrtitus University of Thrace, Department of Civil Engineering, Vas.Sofias 12, Xanthi 67100, Greece*

Abstract

This paper presents a STAMP-based indicator of measuring the inherent, in terms of the system design and development, capability of each system part to provide its agent with Situation Awareness (SA) about the presence of system threats and vulnerabilities that may lead to accidents. An agent is a human or automated controller that possesses reasoning mechanisms and demonstrates a capability to influence others or modify situations. This capability – in as far as it pertains to risk modification - is called "*risk SA provision capability*" (RiskSOAP) and can be modelled in a control loop. This capability is considered as dynamic because it can fluctuate over time due to changes in safety specifications and short- or long-term conditions. In order to demonstrate the fluctuation of the *risk SA provision capability* along the development of an accident, the STAMP-based RiskSOAP indicator is calculated throughout the Überlingen accident timeline. This timeline incorporates four milestones, each one denoting a particular time point in the accident development. The decline in the value of the RiskSOAP indicator is attributed to the presence of flaws and unsafe control actions, through which accident scenarios are verified and the system is headed for an accident. The main conclusion is that in such socio-technical systems there is a tight coupling between the degradation of the *risk SA provision capability* and the degradation of safety.

Peer-review under responsibility of the organizing committee of STAMP EU 2015

Keywords: Safety; STAMP; STPA; EWaSAP; RiskSOAP methodology; timeline

1. Introduction

Complex socio-technical systems require a more holistic reasoning and targeted approaches than those offered by traditional Situation Awareness (SA) models. Current SA measurement techniques are not adequate and/or valid for

* Corresponding author. Tel.: +30-6942-413-863; fax: +30-25410-79678.
E-mail address: mikelachatzimichailidou@gmail.com

Peer-review under responsibility of the organizing committee of STAMP EU 2015
doi:10.1016/j.proeng.2015.11.503

estimating SA in complex socio-technical systems settings [1]. In particular, with the current technological basis (i.e. we cannot constantly monitor human brain functions and reactions to stimuli) it is the cognitive and distributed 'character' of SA in complex socio-technical systems that possibly renders its direct measurement quite a challenging task, if not impossible [1].

In literature there is a debate whether SA is a folk model [2] or a key concept for safety science [3]. This paper, however, does not address that dilemma, but suggests a different approach to SA. We elaborate on an inherent capability of each system part to provide its controller with SA about the presence of system threats and vulnerabilities that may lead to accidents [4]. This inherent capability, called "*risk SA provision capability*" (RiskSOAP), is dynamic given that a system consists of a specific number and type of elements according to its design specifications, thus, the absence and/or inadequate functioning of them may entail the degradation of the *risk SA provision capability*.

If the inadequate or missing system elements are not acknowledged and replaced or fixed, they will probably contribute to a safety drift. That is, the system's defenses will be eroded, particularly in a time of degraded technical system capabilities [5], leaving the system controllers with no or degraded input, thereby rendering them unable to recognize a safety drift and envisage the system's possible future states. Neither a direct measurement nor an assessment of SA shape the primary goal of this research work because the 'measured substance' is different compared to the existing SA measurement techniques. Instead, this work elaborates on the inherent *risk SA provision capability* of the system and demonstrates the relation between the *risk SA provision capability* and safety. On that account, this work presents an alternative approach to SA.

In the Überlingen mid-air collision accident, along with the violated control actions and safety constraints causing a degradation of safety, the official accident investigation reports (e.g. [5]) name specific technical and human services and information content that were lost during the development of the accident, and contributed to it. At the same time, as long as the operative parts of the system were being decreased, the *risk SA provision capability* was also degrading, as this is demonstrated by the gradual decline of the calculated RiskSOAP indicator presented in the remaining of the paper. Thus, there was a negative impact simultaneously on the *risk SA provision capability* and safety due to the 'erosion' of the system's composition. According to the BFU [5] accident investigation report, had the short-term conflict alert (STCA) system not been downgraded, it would have provided the air traffic controller/manager (ATC/ATM) with a visual warning of collision trajectory instead of the auditory alarm alone. Furthermore, the presence of an ATC assistant would have resulted in an even distribution of workload and a more focused attention. These examples of the system losing two elements were identified by the BFU report [5] as systemic causes, and lead, in combination with other factors, to the degradation of the *risk SA provision capability*. In short, the degraded *risk SA provision capability* was a contributing factor in the Überlingen accident, indicating that the RiskSAOP did relate to safety.

Based on the above reasoning, by applying the STAMP-based RiskSOAP indicator throughout the Überlingen accident timeline, the degradation of the *risk SA provision capability* over time is demonstrated. The decline in the value of the RiskSOAP indicator is attributed to the absence or malfunction of specific system elements and their interactions, as well as the presence of flaws and unsafe control actions through which accident scenarios are verified and, in turn, the system is headed for an accident.

2. The RiskSOAP methodology

The RiskSOAP methodology is founded on three already existing approaches, which were combined in a unique manner and executed in the following order (Figure 1): (1) the STAMP Based Process Analysis (STPA) [6]; (2) the Early Warning Sign Analysis based on the STPA (EWaSAP) approach [7], which in conjunction with STPA defines the elements and the characteristics that should be included in the ideal system design; (3) a binary dissimilarity measure to depict the distance between the ideal and the real system design.

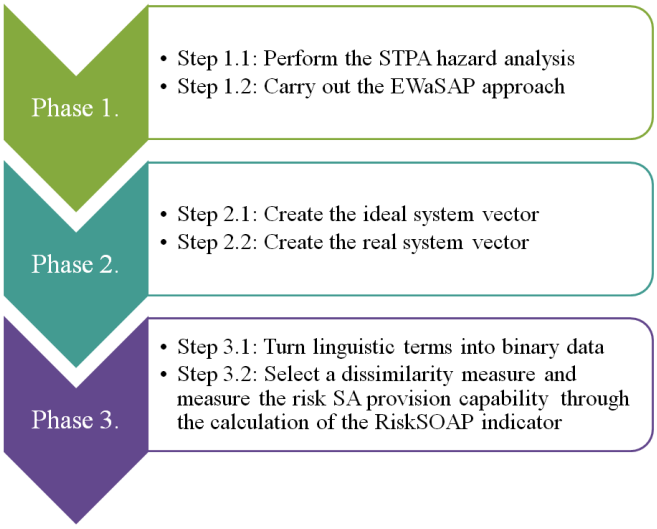

Fig. 1. The phases of the RiskSOAP methodology.

The RiskSOAP methodology can be used either one-off or in an iterative manner. As regards the former, if the aim is to select among alternative systems, then the RiskSOAP can be calculated once, one for each system design. However, in the latter case, when two or more design versions of the same system are about to be compared, then it is feasible to calculate the RiskSOAP indicator as many times as the different alternative versions of the system under consideration [8]. The second case is the one that is applied in this paper.

2.1. STAMP and EWaSAP

Leveson's [9] Systems-Theoretic Accident Model and Processes (STAMP) advocates that accidents represent a complex, dynamic process, meaning that they are not simply chains of component failures. A hazard analysis technique that encapsulates the principles of the STAMP accident causality model is STPA [6]. STPA is a top-down system engineering approach to system safety and can be used early in the system development process to generate high-level safety requirements and constraints. EWaSAP is an add-on to STPA [7] and its aim is to provide a structured method for identifying early warning signs through perceivable sets of data which indicate in a timely manner the presence of flaws and threats to a system [7]. Furthermore, EWaSAP introduces an additional type of control action, the awareness action. An awareness control action allows a controller to provide warning messages and alerts to other controllers inside or outside the system boundary, whenever data indicating the presence of threats or vulnerabilities is received and comprehended.

STPA and EWaSAP could be performed as one process [7] (Table 1), or consecutively by executing STPA first and then EWaSAP.

Table 1. EWaSAP steps as add-ons to STPA

STPA steps and description	EWaSAP steps and description
STPA(1) Identify system hazards & translate them into top-level safety constraints	
	EW(1) Decide if there is anyone outside the system who needs to be informed about the perceived progress of the hazard or about its occurrence
STPA(2a) Create control structure (see Figure 3) STPA(2b) Determine how hazards can occur STPA(2c) Restate inadequate control actions as safety constraints	
	EW(2) Aim: Identify useful sensory services (i.e. video surveillance cameras pointing) installed in or possessed by systems outside of the system in focus and establish synergy EW(2a) For each top level safety constraint identify those signs which indicate its violation EW(2b) Find those systems in the surrounding environment with sensors capable of perceiving the signs defined in EW(2a) & request to establish synergy
STPA(3a) For each element in the control structure create a model of the process it controls STPA(3b) Examine the parts of the control loops to determine if they can contribute to or cause system level hazards	
	EW(3) Aim: Enforce Internal Awareness Actions EW(3a) Describe what needs to be monitored & what type of features/capabilities the sensors must have so that to make the appropriate controllers capable of perceiving: - the signs indicating the occurrence of the flaw - the violation of the assumptions made during the design of the system EW(3b) After design trade-offs and selection of sensors, define which patterns of perceived data indicate the occurrence of the flaw and/or the violation of its designing assumptions EW(3c) Update the process models of the controllers with appropriate awareness and control actions, which should be enforced based on the perceived early warning signs, so that to warn about, adapt to, or eliminate the causal factor to the loss which is present in the system EW(3d) For each perceived warning sign, define its meta-data/attribute values to ensure that it will be perceived and ultimately understood by the appropriate controller/s
STPA(4) Restate any flaws identified as safety constraints & repeat STPA(3a) & STPA(3b)	

2.2. Rogers-Tanimoto Dissimilarity Measure

In the literature there are plenty of distance/dissimilarity measures [10,11], which detect the mismatching bits of two binary data sets. The selection of the proper dissimilarity measure is customised to the assumptions made by the investigator during a specific problem statement.

In this research work Rogers-Tanimoto was chosen as the appropriate dissimilarity measure for comparing the design versions of the same system, on the basis that it is the only dissimilarity measure that gives weight to the

dissimilarities between two compared units by multiplying them by two, i.e. '$2 \times S10$', '$2 \times S01$' [8]. The Rogers-Tanimoto dissimilarity measure is given by the following formula:

$$\mathrm{RTd}(i,r) = \frac{2S10 + 2S01}{S11 + S00 + 2S10 + 2S01} \quad (1)$$

The terms: 'S00', 'S01', 'S10', and 'S11' denote the total number of the corresponding (0,0), (0,1), (1,0), and (1,1) pairs of binary integers, of the two compared units. Figure 2 shows that there is a one-by-one relationship between the binary integers that shape a specific pair.

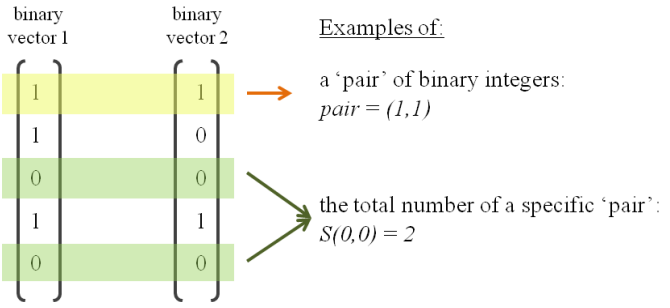

Fig. 2. A graphical explanation of the 'pairs' and 'totals' for the dissimilarity measures.

In dissimilarity measures the following apply:

a. The minimum dissimilarity is '0', meaning that when the dissimilarity of two compared units tends to '1', then the compared units are almost dissimilar.
b. All variables are brought into a common scale, between '0' and '1', i.e. they are normalised.
c. Distance can be defined as a dual of a similarity measure, i.e. $\mathrm{d}(i,r) = 1 - \mathrm{s}(i,r)$. This literally means that a similarity can be expressed as the complementary of the corresponding dissimilarity, and vice versa.

3. Case Study

The RiskSOAP methodology was conducted for the case of the Überlingen mid-air collision accident. In this accident RiskSOAP measures the distance between the 'ideal' system design, as defined by the STPA hazard analysis technique, and the actual system state during different operational phases of the system's life-cycle until the accident. Certain parts of the system were inoperative, meaning that the actual socio-technical system under study was already 'not ideal'; it included flaws and unsafe control actions followed by significant trade-offs and degraded modes of operation. In Figure 3 the elements that were gradually falling apart either from the design or during the operation of the system are depicted with dashed lines.

In *Phase 1*, based on the safety control structure of the systems involved in the Überlingen accident (Figure 3), STPA and EWaSAP were applied for that accident in order to define the 'ideal' system composition. ATC1 and 2 represent the two air traffic controllers being on duty in the ideal case. P1, 2, 3 are the three pilots of the three controlled aircraft and A1, 2, 3 are the three aircraft. TCAS1, 2, 3 are the systems installed in each of the three aircraft.

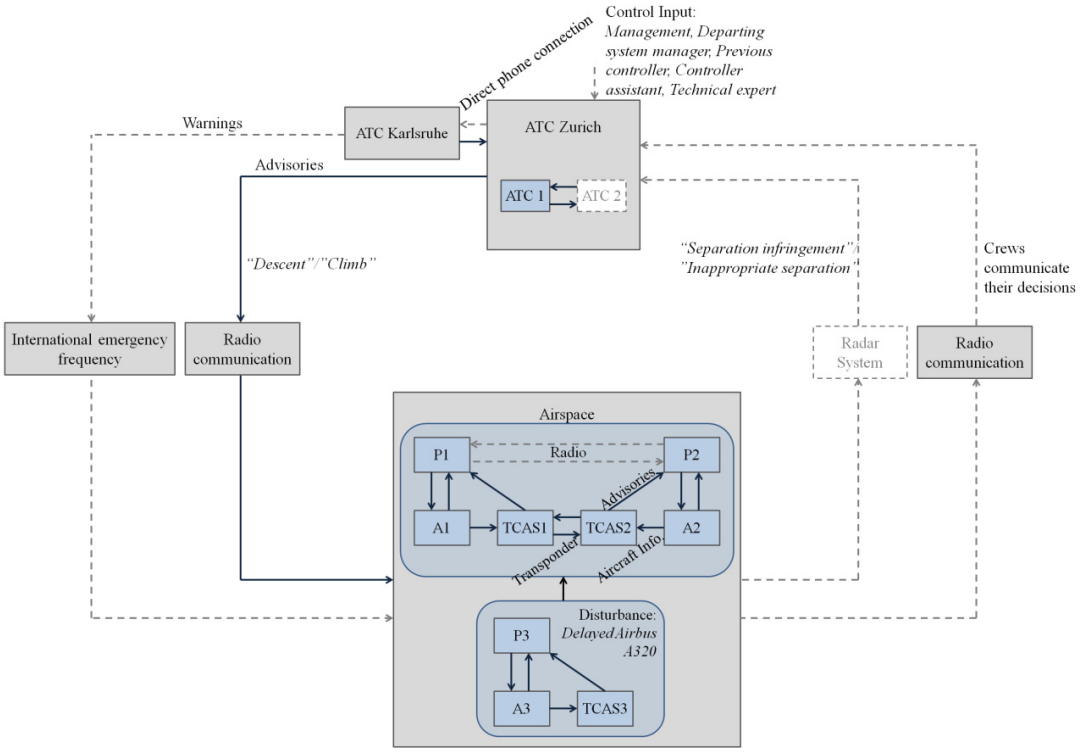

Fig. 3. The safety control structure for the Überlingen case.

The examined accident was: *Loss of human life due to aircraft collision* and the hazard was: *A pair of controlled aircraft violate minimum separation standards.* The total number of safety requirements and sensor characteristics was 279: 119 safety requirements and 152 sensor characteristics were obtained by taking the STPA and EWaSAP steps respectively; 8 mental models and control algorithms resulted from the combination of the responsibilities and safety constraints that each of the controllers of the system should possess. These mental models and control algorithms could be either in the possession of a human or an automated controller depending on the responsibilities of the controller and the process or procedure considered each time.

In *Phase 2*, all safety requirements and sensor characteristics that were derived from the STPA and EWaSAP shaped the ideal system vector, while the values obtained by mapping the present and absent system elements in the real system depicted the real system vector(s).

Since the main purpose of this paper is to demonstrate the degradation of the *risk SA provision capability*, the RiskSOAP methodology and the corresponding indicator are applied throughout the Überlingen accident timeline. This timeline is divided into three sections, i.e. [t-1, t], [t, t+1] and [t+1, t+2], each one defined by a particular starting and ending time point in the accident development. The four time-points (marked on Figure 5) of the Überlingen accident timeline are as follows: “Existing regulations before the collision”, “Beginning of the nightshift”, “Flights become visible on the ATC’s screen”, “Aircraft in a collision trajectory”. It is worth mentioning that at the first time-point, i.e. “Existing regulations before the collision”, the composition of the system incorporated regulations that existed before the collision, as well as any other system element (e.g. human and technical resources, communication channels) communicating the status of the system at that time. The labels assigned to the time-points had to be descriptive in terms of making the differences between the four time-points profound.

In order to demonstrate how the degradation of the *risk SA provision capability* may cause a safety drift in the examined case of the Überlingen accident, the presence/absence of the system elements is mapped four times: (1) at

the time-point 't-1', based on the system's composition according to the regulations before the Überlingen accident occurred; (2) at the time-point 't', when the nightshift begins; (3) at the time-point 't+1', when the conflicting flights become visible on radar; (4) at the time-point 't+2', when the two aircraft are in a collision trajectory. Hence, four vectors were compared to the one that encapsulates the results of STPA and EWaSAP.

In *Phase 3* every component of the 279-sized vector that resulted from the STPA and EWaSAP was equal to '1' because in the examined case of the Überlingen accident it reflected the 'ideal' system design version. As expected, the rest four vectors included both '1' and '0' values due to losses of system elements. Given the binary values assigned to the safety requirements and sensor characteristics, the Rogers-Tanimoto dissimilarity measure was calculated. The precise values of the terms of the Rogers-Tanimoto formula [equation (1)] are given in Table 2. With the use of the Rogers-Tanimoto measure, the value of the RiskSOAP indicator was calculated four times; one for each of the four time points shown in Table 2. The overall numerical results of the analysis are given in Table 2. An indicative illustration of how the vectors are compared to each other with the used of the dissimilarity measure is given in Figure 4.

Table 2: The degradation of the *risk SA provision capability* given in numbers.

STPA & EWaSAP	Four milestones of the Überlingen accident timeline			
	♦ t-1	♦ t	♦ t+1	♦ t+2
Present 279	74	65↓	63↓	50↓
Absent - ('ideal')	205	214↑	216↑	229↑
RiskSOAP indicator via RTd(i,r)	=0.8471	=0.8682↑	=0.8727↑	=0.9016↑

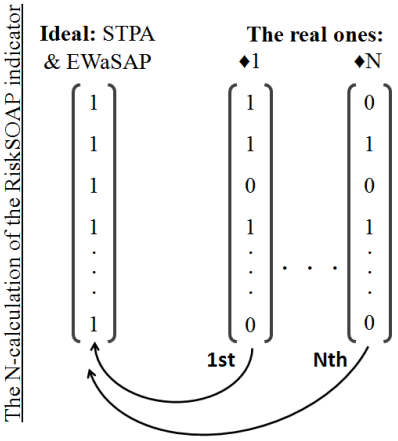

Fig. 4. Calculation of the Rogers-Tanimoto measure and RiskSOAP indicator for N real vectors compared to the ideal one

According to Table 2, the number of system elements being present in the system composition diminishes, i.e. '↓', while the number of those being absent expands, i.e. '↑', along the accident's course. Furthermore, the last row of Table 2 reveals that the decline, i.e. '↑', of the RiskSOAP indicator evolves parallel to the accident development timeline. That is, every time a technical (e.g., main radar system), or a human (e.g., second ATC) service is lost, a further increase in the value of the RiskSOAP indicator is noticed. The gradual decline of the value signifies a further degradation of the *risk SA provision capability* of the system and a further deterioration of system safety.

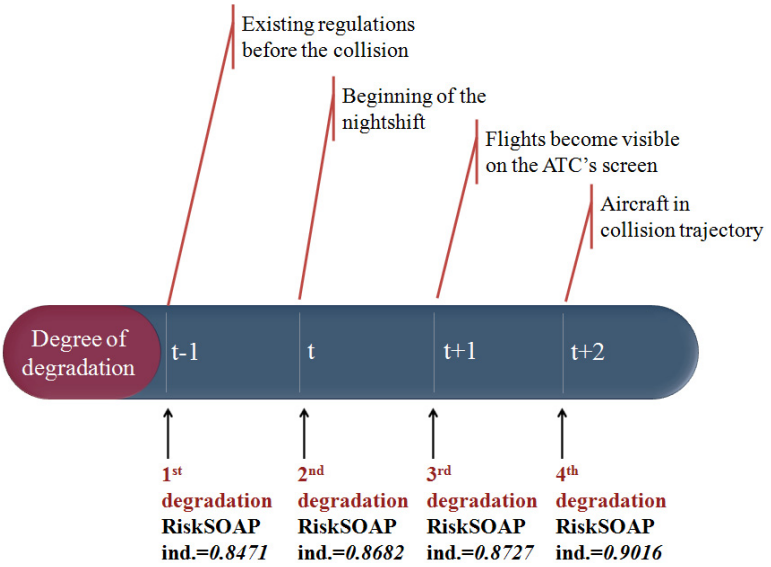

Fig. 5. The milestone events after which the *risk SA provision capability* was measured.

4. Discussion

Aside from the degradation of the *risk SA provision capability* that did happen parallel to the development of the Überlingen accident, the decrease of the RiskSOAP indicator is attributed to the presence of flaws and unsafe control actions, through which accident scenarios are verified and the system is headed for an accident.

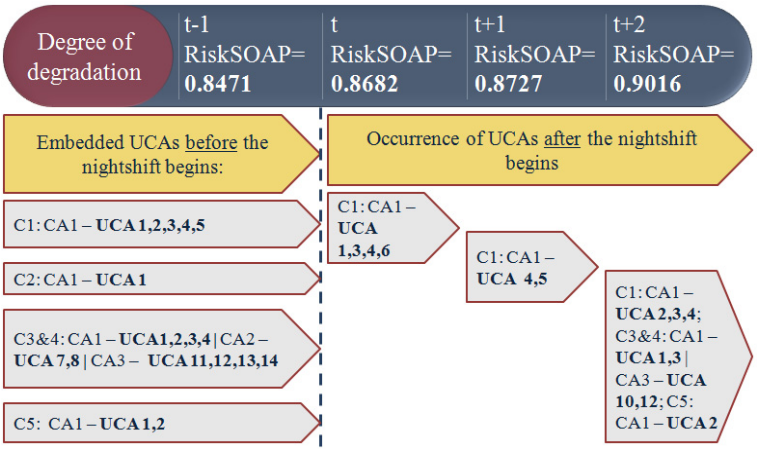

Fig. 6. Accident scenarios verified when he *risk SA provision capability* was degrading.

Figure 6 illustrates the gradual degradation of the *risk SA provision capability* for the systems involved in the Überlingen accident along with the unsafe control actions (UCAs) that were either embedded in the system, prior to

the 'fatal' nightshift, or occurred during that nightshift. In either of the two cases, the UCAs were attributed to the presence of specific flaws that, according to STPA, could have led to accident(s), as it did happen in reality.

Ideally, according to the STPA hazard analysis, none of these UCAs should have occurred. In Figure 6, every 'Cx' represents one of the controllers (i.e. C1: ATC Zurich, C2: ATC Karlsruhe, C3&4: aircraft crews, C5: TCAS) of the system each time. For every controller, Figure 6 displays the corresponding control actions (CAs), e.g. C1:CA1. The displayed CAs were, in fact, the inadequate and unsafe control actions, corresponding to a specific controller and a specific CA he/she/ it was in charge to enforce (e.g. C1: CA1 - UCA1,2,3,4,5). Figure 6 includes the UCAs, i.e. CA (a) not provided, (b) provided at the wrong time, or (c) provided but not followed. that were involved in the Überlingen accident. The identified flaws were the causal factors considered to create the hazardous scenarios and contribute to the degradation of the *risk SA provision capability*. The UCAs given in Figure 6 were created by the above-mentioned flaws, which were either inherent in the system prior to the accident or induced during the migration of systems toward states of increasing risk [6]. It would be useful to note that there were UCAs being stopped at some time during the accident development, e.g. C2: CA1 - UCA1, and others being applied throughout the whole accident, like the C1: CA1 - UCA1.

The results suggest that there is a relation between safety and the *risk SA provision capability* of a socio-technical system. Due to space-saving reasons, some indicative examples of the unsafe control actions identified by applying the STPA hazard analysis are displayed in Table 3.

Table 3: Unsafe Control Actions defined for the Überlingen accident.

Safety requirements & sensor characteristics not met (flaws)		UCAs/ Accident scenarios	Milestones
1	Air navigation service companies should not tolerate one-manned operations	Separate two aircraft provided (by the ATC) too late when two aircraft are too close to each other to start maneuvering and avoid collision (C1*-CA1-UCA3)	t-1
2	The Bypass System should be always available to the ATC, or in cases where it is out of service the ATC should be informed	Warning not provided to the ATC Zurich in case when he does not realise the collision trajectory (C2*-CA1-UCA1)	
3	National civil aviation organisation should not be affected by national culture	"Fly according to OP and FP" provided when the two crews do not adhere to the same standardised procedures (C3&4*-CA1-UCA2)	
4	Additional features should be added to the ATC displays after identified incidents or changes in practices	Separate two aircraft not provided (by the TCAS, traffic alert and collision avoidance system) when aircraft in collision trajectory (C5*-CA1-UCA1)	
5	There should be a downlink in place to pass the TCAS advisories to the ATC	Separate two aircraft provided when TCAS issues opposite advisory (compared to the ATC) (C5*-CA1-UCA2)	
6	Automated systems or audits should provide necessary error checking to detect ATC's possible errors	Separate two aircraft not provided (by the ATC) when two aircraft in collision trajectory (C1*-CA1-UCA1)	t
7	A sensor should be able to measure the (high) traffic	Separate two aircraft provided wrongly: pair-wise advisories issued to the two crews are not complementary to each other; conflicting conditions emerge (C1*-CA1-UCA4)	
8	A sensor should detect whether the two aircraft have violated the minimum separation threshold	Separate two aircraft provided wrongly: conflicting advisories between ATC and TCAS when it is not clear for the crew(s) on which one to adhere to (C1*-CA1-UCA5)	t+1
9	The ATC should be aware that the TCAS has the highest priority as a collision avoidance controlling tool	Separate two aircraft provided when TCAS issues opposite advisory (compared to the ATC) (C1*-CA1-UCA2)	t+2
10	The crew should not ignore the copilot when he communicates a crucial information	Adhere to TCAS provided too late when there is not much time left for maneuvers; collision avoidance not ensured (C3&4*-CA3-UCA12)	
11	The crew(s) should verbally acknowledge the ATC advisory and/or the instructions given by the TCAS	Separate two aircraft provided when TCAS issues opposite advisory (compared to the ATC) (C5*-CA1-UCA2)	
12	A sensor should calculate the relative location of the two aircraft in a timely manner	Separate two aircraft provided too late (by the ATC) when two aircraft are too close to each other to start maneuvering and avoid collision (C1*-CA1-UCA3)	

The safety requirements & sensor characteristics listed in column 2 were all absent (i.e. 0) from the system involved in the accident. In an attempt to indicate the point on the overall safety control structure at which those elements were

absent, column 3 assigns each of them to the controller affected by their absence. In this way, the binary integers (i.e. 0 and 1) were overlaid onto the hierarchical safety control structure and the hierarchy of the system as well.

5. Conclusion

In the Überlingen mid-air collision accident examined herein the *risk SA provision capability* was degrading gradually and in parallel to the degradation or loss of technical and human services, as well as information. However, according to the BFU [5] official accident investigation report, the deterioration or loss of those system elements along with the violated control actions and safety constraints caused a safety drift and thus contributed to the accident. This paper provided evidence of the relation between the *risk SA provision capability* and safety. It was also found that the degraded *risk SA provision capability* was a contributing factor in the Überlingen accident. Applying, therefore, the STAMP-based RiskSOAP indicator throughout the Überlingen accident timeline, the degradation of the *risk SA provision capability* was demonstrated and given a quantitative description.

The results obtained for the Überlingen case suggested that the degradation of the *risk SA provision capability* over time, as clearly revealed by the gradual decline of the RiskSOAP indicator, was aligned with the accident timeline. Thus, the interpretation of the RiskSOAP indicator can lead to conclusions about the degradation or enhancement of the situation that the system is engaged in. Another interesting finding was that every time the value of the RiskSOAP indicator was calculated, the accident scenarios, which actually lead to the specific accident, were verified.

More elaborate and extended experiments will further support the findings of this research. Moreover, the relation between safety and the *risk SA provision capability* of a system can be cross-checked through the application of the RiskSOAP methodology to additional engineering applications and to other research fields, such as healthcare, as well.

References

[1] M.M. Chatzimichailidou, A. Protopapas, I.M. Dokas, Seven issues on distributed situation awareness measurement in complex aocio-technical systems, Complex Systems Design & Management (2015) 105–117.
[2] S.W.A. Dekker, The danger of losing situation awareness, Cognition, Technology & Work 17 (2015) 159–161.
[3] P.M. Salmon, G.H. Walker, N.A. Stanton, Broken components versus broken systems: why it is systems not people that lose situation awareness, Cognition, Technology & Work 17 (2015) 179–183.
[4] M.M. Chatzimichailidou, N.A. Stanton, I.M. Dokas, The concept of risk situation awareness provision: towards a new approach for assessing the DSA about the threats and vulnerabilities of complex socio-technical systems, Safety Science 79 (2015) 126–138.
[5] German Federal Bureau of Aircraft Accident Investigation, BFU Überlingen Investigation Report - Reference AX001-1-2/02, Bundesstelle für Flugunfalluntersuchung, Braunschweig, 2004.
[6] N. Leveson, Engineering A Safer World: Systems Thinking Applied to Safety, MIT Press, Cambridge, MA, 2011.
[7] I.M. Dokas, J. Feehan, S. Imran, EWaSAP: an early warning sign identification approach based on a systemic hazard analysis, Safety Science 58 (2013) 11–26.
[8] M.M. Chatzimichailidou, I.M. Dokas, Introducing RiskSOAP to communicate the distributed situation awareness of a system about safety issues: An application to a robotic system, Ergonomics (2015) 1–14.
[9] N. Leveson, A new accident model for engineering safer systems, Safety Science 42(4) (2004) 237–270.
[10] B. Zhang, S.N. Srihari, Properties of binary vector dissimilarity measures, in: Computer Vision, Pattern Recognition, and Image Processing, Springer, 2003.
[11] S.S. Choi, S.H. Cha, C.C. Tappert, A survey of binary similarity and distance measures, Journal of Systemics, Cybernetics and Informatics 8(1) (2010) 43–48.